

## DISK CONTROLLER, STORAGE SYSTEM, AND METHOD FOR CONTROLLING THE SAME

Patent Number: JP2003345515  
Publication date: 2003-12-05  
Inventor(s): FUJIMOTO TAKEO; HONMA HISAO  
Applicant(s): HITACHI LTD  
Requested Patent: JP2003345515  
Application: JP20020152864 20020527  
Priority Number(s):  
IPC Classification: G06F3/06; G06F12/16; G06F13/12;  
EC Classification:  
Equivalents:

### Abstract

**PROBLEM TO BE SOLVED:** To achieve high speed data transfer by performing access from a data transfer adapter and a controller to a data transfer parameter.

**SOLUTION:** This disk controller is provided with a cache memory for temporarily storing input/output data to a disk device in which data are stored, a channel adapter part having a processor for controlling the transfer of data between a host and the cache memory and a data transfer adapter means, and a disk adapter part having a processor for controlling the transfer of data between the disk device and the cache memory and the data transfer adapter means. The processor of the channel adapter part or the disk adapter part stores a parameter for controlling the transfer of data in a memory in the processor, and the data transfer adapter means of the channel adapter part or the disk adapter part reads the parameter stored in the memory, and executes the transfer of data based on the read parameter.

Data supplied from theesp@cenet database - I2

(19) 日本国特許庁 (JP)

## (12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-345515

(P2003-345515A)

(43) 公開日 平成15年12月5日 (2003.12.5)

(51) Int.Cl.<sup>1</sup>  
 G 0 6 F 3/06  
 3 0 1  
 3 0 2  
 12/16 3 1 0  
 13/12 3 3 0  
 13/28 3 1 0

識別記号  
 3 0 1  
 3 0 2  
 3 1 0  
 3 3 0  
 3 1 0

F I  
 G 0 6 F 3/06  
 3 0 1 M 5 B 0 1 4  
 3 0 2 A 5 B 0 1 8  
 12/16 3 1 0 R 5 B 0 6 1  
 13/12 3 3 0 A 5 B 0 6 5  
 13/28 3 1 0 E

マークト<sup>1</sup> (参考)

審査請求 未請求 請求項の数12 O L (全 7 頁)

(21) 出願番号 特願2002-152864(P2002-152864)

(22) 出願日 平成14年5月27日 (2002.5.27)

(71) 出願人 000005108

株式会社日立製作所  
東京都千代田区神田駿河台四丁目6番地

(72) 発明者 藤本 健雄

神奈川県小田原市中里322番地2号 株式  
会社日立製作所 R A I D システム事業部内

(72) 発明者 本間 久雄

神奈川県小田原市中里322番地2号 株式  
会社日立製作所 R A I D システム事業部内

(74) 代理人 100071283

弁理士 一色 健輔 (外4名)

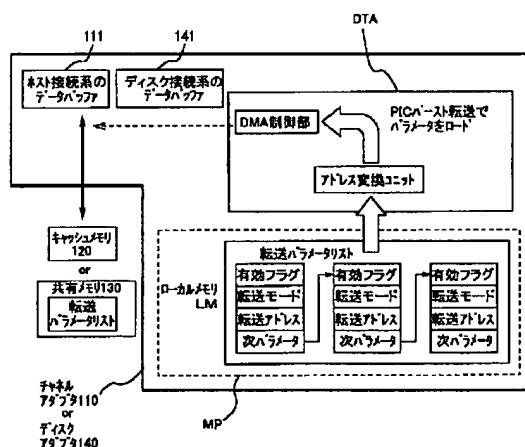
最終頁に続く

(54) 【発明の名称】 ディスク制御装置、ストレージシステム、及びその制御方法

(57) 【要約】 (修正有)

【課題】 データ転送パラメータをデータ転送アダプタおよび制御装置からアクセスできるようにし高速なデータ転送を可能とする。

【解決手段】 データを記憶するディスク装置への入出力データを一時的に格納するキャッシュメモリと、ホストとキャッシュメモリとの間のデータの転送を制御するプロセッサ及びデータ転送アダプタ手段を有するチャネルアダプタ部と、ディスク装置とキャッシュメモリとの間のデータの転送を制御するプロセッサ及びデータ転送アダプタ手段を有するディスクアダプタ部とを備えたのであって、チャネルアダプタ部又はディスクアダプタ部のプロセッサは、データの転送を制御するためのパラメータをプロセッサ内のメモリに格納し、チャネルアダプタ部又はディスクアダプタ部のデータ転送アダプタ手段は、メモリに格納されたパラメータを読み出し、読み出したパラメータに基づいてデータの転送を実行することとする。



(2)

特開2003-345515

2

## 【特許請求の範囲】

【請求項1】 データを記憶するディスク装置への入出力データを一時的に格納するキャッシュメモリと、ホストと前記キャッシュメモリとの間のデータの転送を制御するプロセッサ及びデータ転送アダプタ手段を有するチャネルアダプタ部と、前記ディスク装置と前記キャッシュメモリとの間のデータの転送を制御するプロセッサ及びデータ転送アダプタ手段を有するディスクアダプタ部とを備えたディスク制御装置において、

前記チャネルアダプタ部又は前記ディスクアダプタ部の前記プロセッサは、前記データの転送を制御するためのバラメータを前記プロセッサ内のメモリに格納し、前記チャネルアダプタ部又は前記ディスクアダプタ部の前記データ転送アダプタ手段は、前記メモリに格納された前記バラメータを読み出し、読み出した前記バラメータに基づいて前記データの転送を実行することを特徴とするディスク制御装置。

【請求項2】 前記プロセッサは、前記メモリに格納される前記バラメータのアドレスを変更可能であることを特徴とするディスク制御装置。

【請求項3】 前記プロセッサ内の前記メモリに障害が発生した場合には、前記バラメータを該メモリにおける別のバンクに格納すべく、前記プロセッサは、該バラメータのアドレスを変更することを特徴とする請求項2に記載のディスク制御装置。

【請求項4】 前記チャネルアダプタ部及び前記ディスクアダプタ部に接続された共有メモリを更に備え、前記プロセッサに障害が発生した場合には、前記バラメータを前記共有メモリに格納すべく、該バラメータのアドレスを前記共有メモリに変更し、他の正常な前記プロセッサが、前記バラメータを前記共有メモリに格納することを特徴とする請求項2に記載のディスク制御装置。

【請求項5】 データを記憶するディスク装置と、前記ディスク装置への入出力データを一時的に格納するキャッシュメモリと、ホストと前記キャッシュメモリとの間のデータの転送を制御するプロセッサ及びデータ転送アダプタ手段を有するチャネルアダプタ部と、前記ディスク装置と前記キャッシュメモリとの間のデータの転送を制御するプロセッサ及びデータ転送アダプタ手段を有するディスクアダプタ部とを備えたディスク制御装置において、

前記チャネルアダプタ部又は前記ディスクアダプタ部の前記プロセッサは、前記データの転送を制御するためのバラメータを前記プロセッサ内のメモリに格納し、前記チャネルアダプタ部又は前記ディスクアダプタ部の前記データ転送アダプタ手段は、前記メモリに格納された前記バラメータを読み出し、読み出した前記バラメータに基づいて前記データの転送を実行することを特徴とするストレージシステム。

【請求項6】 前記プロセッサは、前記メモリに格納される前記バラメータのアドレスを変更可能であることを特徴とするストレージシステム。

【請求項7】 前記プロセッサ内の前記メモリに障害が発生した場合には、前記バラメータを該メモリにおける別のバンクに格納すべく、該バラメータのアドレスを変更することを特徴とする請求項6に記載のストレージシステム。

【請求項8】 前記チャネルアダプタ部及び前記ディスクアダプタ部に接続された共有メモリを更に備え、前記プロセッサに障害が発生した場合には、前記バラメータを前記共有メモリに格納すべく、該バラメータのアドレスを前記共有メモリに変更し、他の正常な前記プロセッサが、前記バラメータを前記共有メモリに格納することを特徴とする請求項6に記載のストレージシステム。

【請求項9】 データを記憶するディスク装置への入出力データを一時的に格納するキャッシュメモリと、ホストと前記キャッシュメモリとの間のデータの転送を制御するプロセッサ及びデータ転送アダプタ手段を有するチャネルアダプタ部と、前記ディスク装置と前記キャッシュメモリとの間のデータの転送を制御するプロセッサ及びデータ転送アダプタ手段を有するディスクアダプタ部とを備えたディスク制御装置の制御方法であって、

前記チャネルアダプタ部又は前記ディスクアダプタ部の前記プロセッサは、前記データの転送を制御するためのバラメータを前記プロセッサ内のメモリに格納し、前記チャネルアダプタ部又は前記ディスクアダプタ部の前記データ転送アダプタ手段は、前記メモリに格納された前記バラメータを読み出し、読み出した前記バラメータに基づいて前記データの転送を実行することを特徴とするディスク制御装置の制御方法。

【請求項10】 前記プロセッサは、前記メモリに格納される前記バラメータのアドレスを変更可能であることを特徴とするディスク制御装置の制御方法。

【請求項11】 前記プロセッサ内の前記メモリに障害が発生した場合には、前記バラメータを該メモリにおける別のバンクに格納すべく、該バラメータのアドレスを変更することを特徴とする請求項10に記載のディスク制御装置の制御方法。

【請求項12】 前記ディスク制御装置は、前記チャネルアダプタ部及び前記ディスクアダプタ部に接続された共有メモリを更に備え、

前記プロセッサに障害が発生した場合には、前記バラメータを前記共有メモリに格納すべく、該バラメータのアドレスを前記共有メモリに変更し、他の正常な前記プロセッサが、前記バラメータを前記共有メモリに格納することを特徴とする請求項10に記載のディスク制御装置の制御方法。

50 【発明の詳細な説明】

(3)

特開2003-345515

3

## 【0001】

【発明の属する技術分野】この発明は、ディスク制御装置、ストレージシステム、及びその制御方法に関する。

## 【0002】

【従来の技術】ディスクシステムは、図6に示すように、ディスク制御装置100と、ディスクアレイ200を備える。ディスク制御装置100は、複数のチャネルアダプタ110と、キャッシュメモリ120と、共有メモリ130と、複数のディスクアダプタ140とを備える。ディスクアレイ200は、データを記憶する複数のディスク装置210で構成される。ここでは、便宜上、キャッシュメモリ120と共有メモリ130とを同一のブロックで示しているが別個のハードウェアで構成される。

【0003】チャネルアダプタ110は、プロセッサ(CPUを中心としてローカルメモリ等の周回回路を含む処理装置)MPとデータ転送アダプタDTAを備える。このチャネルアダプタ110は、キャッシュメモリ120及び共有メモリ130とホスト300とに接続され、キャッシュメモリ120とホスト300との間のデータ転送を制御する。キャッシュメモリ120は、出入力データを一時的に格納するバッファの役割をも果たす。共有メモリ130は、キャッシュメモリ120の管理用ディレクトリ情報や、チャネルアダプタ110及びディスクアダプタ140の各プロセッサMP間の通信情報等を含む制御情報が格納される。

【0004】ディスクアダプタ140は、プロセッサMPとデータ転送アダプタDTAを備える。このディスクアダプタ140は、キャッシュメモリ120及び共有メモリ130とディスク装置210とに接続され、キャッシュメモリ120とディスク装置210との間のデータ転送を制御する。

【0005】チャネルアダプタ110及びディスクアダプタ140における各データ転送アダプタDTAは、データ転送を司る機能を有し、レジスタ群及びDMA機構を備える。レジスタ群は、多数のレジスタで構成される。これらレジスタに、転送モード(書き込みや読み出し等)や転送アドレス及び転送データ長等のデータ転送に必要なパラメータ(以下、「転送パラメータ」と称する)が格納される。DMA機構は、キャッシュメモリ120へのアクセス用チャネルであり、レジスタから読み出したデータ転送パラメータに基づき、キャッシュ120とのデータ転送を制御する。プロセッサMPは、データ転送アダプタDTAの各レジスタに対し、転送パラメータを設定する。

## 【0006】

【発明が解決しようとする課題】近年の技術進歩により、プロセッサMPの性能が向上している。このため、プロセッサMPが転送パラメータを設定するにあたり、プロセッサMP自身の内部にあるローカルメモリへア

4

セスする速度が向上している。このローカルメモリへのアクセスに比べると、データ転送アダプタDTAの各レジスタに対してアクセスする場合の方が、時間が長くかかる結果となる。

【0007】一方、転送するデータ量が飛躍的に増大する傾向にある現在においては、オーバヘッドを隠すべく、多様な転送パラメータを用いているものの、データ転送アダプタDTAのレジスタに対するアクセスも増えている。プロセッサMPの性能向上にも拘わらず、やはりプロセッサMPによる転送パラメータの設定時間が問題となり、データ転送の性能問題に発展する。

【0008】また他に、プロセッサMPが破損すると、データ転送アダプタDTAを通じてデータを転送することができなくなるといった問題がある。

【0009】本発明は、このような課題に鑑みてなされたもので、ディスク制御装置、ストレージシステム、及びその制御方法を提供することを目的とする。

## 【0010】

【課題を解決するための手段】データを記憶するディスク装置への入出力データを一時的に格納するキャッシュメモリと、ホストと前記キャッシュメモリとの間のデータの転送を制御するプロセッサ及びデータ転送アダプタ手段を有するチャネルアダプタ部と、前記ディスク装置と前記キャッシュメモリとの間のデータの転送を制御するプロセッサ及びデータ転送アダプタ手段を有するディスクアダプタ部とを備えたのであって、前記チャネルアダプタ部又は前記ディスクアダプタ部の前記プロセッサは、前記データの転送を制御するためのパラメータを前記プロセッサ内のメモリに格納し、前記チャネルアダプタ部又は前記ディスクアダプタ部の前記データ転送アダプタ手段は、前記メモリに格納された前記パラメータを読み出し、読み出した前記パラメータに基づいて前記データの転送を実行することとする。

【0011】その他、本願が開示する課題、及びその解決方法は、発明の実施形態の欄及び図面により明らかにされる。

## 【0012】

【発明の実施の形態】本明細書の記載により、少なくとも次のことが明らかにされる。前記プロセッサは、前記メモリに格納される前記パラメータのアドレスを変更可能であることとしてもよい。

【0013】また、前記プロセッサ内の前記メモリに障害が発生した場合には、前記パラメータを該メモリにおける別のバンクに格納すべく、該パラメータのアドレスを変更することとしてもよい。

【0014】さらに、前記ディスク制御装置は、前記チャネルアダプタ部及び前記ディスクアダプタ部に接続された共有メモリを更に備え、前記プロセッサに障害が発生した場合には、前記パラメータを前記共有メモリに格納すべく、該パラメータのアドレスを前記共有メモリに

50

(4)

特開2003-345515

5

変更し、他の正常な前記プロセッサが、前記パラメータを前記共有メモリに格納することとしてもよい。

【0015】この発明を実現化できた背景として、従来の技術では、プロセッサMPは、チャネルアダプタ側やディスクアダプタ側のデータ転送アダプタDTAの各レジスタに対し、転送パラメータを格納していた。これに対し、本実施の形態にあっては、この転送パラメータの格納先をプロセッサMP内部のローカルメモリ又は共有メモリとした。

【0016】つまり、プロセッサMPは、わざわざ、その外部のデータ転送アダプタDTA側のレジスタに転送パラメータを設定していた。その理由としては、データ転送アダプタDTAは、設定された転送パラメータに基づいてデータ転送を行う程度の機能しか持ち得なかったからである。

【0017】一方、従来の技術に対する主たる相違点として、本実施の形態では、データ転送アダプタDTAがアドレス変換ユニットを有することでPCI(Peripheral Component Interconnect)バスのマスタ機能を実現し、インテリジェント化を図っている。このPCIバスのマスタ機能により、データ転送アダプタDTAは、転送パラメータをローカルメモリ又は共有メモリから能動的に読み出すことができる。

【0018】したがって、プロセッサMPは、転送パラメータの格納先を、データ転送アダプタDTA側のレジスタではなく、プロセッサMP内部のローカルメモリ又は共有メモリとすることができます。このため、プロセッサMPにとっては、転送パラメータの設定の高速化が図れ、その分、他の演算や論理等の他の処理に集中できる。つまり、プロセッサMPの負荷分散により、プロセッサMPが行う処理全般に関する性能を向上させることができる。

【0019】本実施の形態に関し、以上説明した従来の技術に対する相違点以外は、ディスクシステムの全体的な構成も含め、従来の技術の欄で図6を用いて説明した事項と共通する。よって、本実施の形態の特徴を示す相違点を中心にして、本発明の実施例を説明する。

【0020】

【実施例】本発明に係る実施例につき、図面を参照して説明する。本実施の形態の特徴を端的に示す機能ブロック図を図1に示す。マイクロプロセッサ(プロセッサ)MPのローカルメモリ(メモリ)LMには、転送パラメータのリストが格納される。通常、ローカルメモリLMにおいて、マイクロプロセッサ(プロセッサ)MPのCPUから最も近いパンクに格納する。また、マイクロプロセッサMPが故障した場合には、共有メモリ130に転送パラメータのリストが格納される。一方、データ転送アダプタ(データ転送アダプタ手段)DTAは、DMA制御部(図6の「DMA機構」と同じ)の他にアドレス変換ユニットを備える。このアドレス変換ユニット

6

は、アドレス変換機能を実現する論理回路を用いて構成される。また、マイクロプロセッサのハードウエアとしては、公知のものが用いられ、例えば図2に示す良く知られた構成のものが用いられる。

【0021】まず、転送パラメータのリストについて説明する。このリストは、一セット又は複数セットの転送パラメータで構成される。一セットあたりの転送パラメータには、有効フラグ、転送モード、転送アドレス及び次パラメータが含まれる。有効フラグは転送パラメータ

10 自身が有効か否かを示し、転送モードは、データ転送がリード、ライトあるいはキャッシュメモリ120内のコピーのいずれかを示す。転送アドレスとは、キャッシュメモリ120の空間、チャネルアダプタ(チャネルアダプタ部)110のデータバッファ111、あるいはディスクアダプタ(ディスクアダプタ部)140のデータバッファ141における転送先アドレス及び転送元アドレスである。次パラメータとは、転送パラメータが複数セットがある場合、次に読み出すセットが格納された転送パラメータの開始アドレスである。

20 【0022】次に、データ転送アダプタDTAのアドレス変換ユニットによるPCIバスのマスタ機能について説明する。このPCIバスのマスタ機能によって、マイクロプロセッサMPのローカルメモリ120あるいは共有メモリ130における特定の領域をPCIのI/Oアドレス空間にマッピングできる。ここで、マッピングとは、マイクロプロセッサMPのCPUが認識可能なローカルメモリLMあるいは共有メモリ130のアドレスをPCIのI/Oアドレス空間のアドレスに対応付けて変換する動作を指す。以後、この動作をマッピングと表現する。チャネルアダプタ110及びディスクアダプタ140の各マイクロプロセッサMPは、これらを含むディスク制御装置が起動した際の初期設定時に、ローカルメモリLMにおける特定のアドレスに対応付けてデータ転送アダプタDTA側にアドレスを設定しておく。

30 【0023】そして、図3のフローチャートに示すように、マイクロプロセッサMPが自己のローカルメモリ(メモリ)LMに転送パラメータのリストが格納された(S10: YES)、アドレス変換ユニットは、初期設定時に設定されたアドレスをPCIのI/Oアドレス空間にマッピングする(S20)。次に、アドレス変換ユニットは、マッピングされたアドレスに基づき、ローカルメモリの転送パラメータのリストをPCIのバースト転送モードで自動的に読み出す(S30)。

40 【0024】このようにして読み出された転送パラメータのリストに基づき、DMA制御部によりデータ転送が実行される(S40)。先ず、データ転送アダプタDTAは、転送パラメータの有効フラグが「ON」であれば、転送モード及び転送アドレスに従い、データ転送を開始する。このデータ転送が正常に終了すると、次パラメータの格納アドレスが有効に設定されているか否かを

(5)

特開2003-345515

7

8

確認する。この格納アドレスが有効に設定されている場合には、その開始アドレスをPCIのI/O空間にマッピングし、次の転送バラメータに基づくデータ転送を実行する。

【0025】次に、図1に加えて従来の技術での説明を利用した図6も流用し、データをディスク装置210から読み出して上位のホスト300に転送する例を説明する。ホスト300のリード命令を受けた一つのチャネルアダプタ110は、まず共有メモリ130に格納されたキャッシュメモリ120に関する管理情報を参照し、リード対象のデータがキャッシュメモリ120に格納されているかを確認する。この確認の結果、キャッシュメモリ120に存在していれば、キャッシュメモリ120からデータを読み出してホスト300へ転送する。反対に、キャッシュメモリ120にデータが存在していない場合には、共有メモリ130を介し、チャネルアダプタ110及びディスクアダプタ140の各マイクロプロセッサMP間の通信を行い、ディスクアダプタ140へディスクアレイ200からのデータの読み出しを要求する。この要求を受けたディスクアダプタ140は、リードの対象となるデータの格納場所を演算し、該当するデータをキャッシュメモリ120に転送する。チャネルアダプタ110は、共有メモリ130を参照し、確立されたデータブロックについて、キャッシュメモリ120からホスト300に転送する。

【0026】ここで、データ転送が異常に終了した場合には、データ転送アダプタDTAの割り込み信号によりデータ転送の処理を中止し、この異常終了をマイクロプロセッサMPに通知する。

【0027】ところで、ローカルメモリLMに障害が発生した場合(S100: YES)には、図4のフローチャートに示すように、所定の割り込み信号でマイクロプロセッサMPに通知する。そして、マイクロプロセッサMPは、PCIのI/O空間にマッピングするアドレスを変更する(S110)。ことにより、ローカルメモリLMが備える複数のバンクのうち正常なバンクに転送バラメータのリストの格納をする(S120)。

【0028】また、マイクロプロセッサMPに障害が発生して閉塞した場合を想定する。この場合、図5のフローチャートに示すように、データ転送アダプタDTA側で、マイクロプロセッサMPの障害検出を契機として、PCIのI/O空間にマッピングするアドレスを共有メモリ130に自動的に変更する(S200→S210)。そして、共有メモリ130に転送バラメータのリストが格納されると、この共有メモリ130に転送バラメータのリストに基づきデータ転送を実行する(S220)。このことにより、故障中のマイクロプロセッサMPの障害が修復されるまでの間、このマイクロプロセッサMPを含むチャネルアダプタ110及びディスクアダプタ140を通じたデータ転送については、他のチャネルアダプタ110及びディスクアダプタ140を通じたデータ転送については、他のチャネルアダプタ110及びディスクアダプタ140の正常なマイクロプロセッサMPでもって代行させることができる。

【0029】本実施の形態にあっては、次の効果を奏する。プロセッサは、転送バラメータの格納先を、データ転送アダプタ側のレジスタではなく、プロセッサ内部のローカルメモリ又は共有メモリとすることができます。このため、プロセッサにとっては、転送バラメータの設定の高速化が図れ、その分、他の演算や論理等の他の処理に集中できる。つまり、プロセッサの負荷分散により、プロセッサが行う処理全般に関する性能を向上させることができる。

【0030】また、プロセッサが行う転送バラメータの設定とは独立して非同期に、データ転送アダプタが主体的にデータ転送を司ることができる。

【0031】また、転送バラメータを設定するローカルメモリに故障が発生しても、他のメモリバンクに転送バラメータの格納先を変更する。このことで、データ転送を1継続できる。さらに、プロセッサが故障しても、転送バラメータの格納先を共用メモリに変更する。このことで、他のチャネルアダプタ及びディスクアダプタの正常なマイクロプロセッサがデータ伝送を代行して継続できる。したがって、障害発生時のデータ保全性を高めることができる。

【0032】

【発明の効果】プロセッサによるデータ転送のバラメータ設定の高速化が図れる。

【図面の簡単な説明】

【図1】 本発明の一実施の形態に係るディスク制御装置の特徴を端的に示す機能ブロック図である。

【図2】 本発明の一実施の形態に係るマイクロプロセッサのハードウェアの構成例を示すブロック図である。

【図3】 本発明の一実施の形態に係るデータの転送処理を示すフローチャートである。

【図4】 本発明の一実施の形態に係るローカルメモリLMに障害が発生した場合の対策を示すフローチャートである。

【図5】 本発明の一実施の形態に係るマイクロプロセッサMPに障害が発生した場合の対策を示すフローチャートである。

【図6】 従来のディスク制御装置のブロック図である。

【符号の説明】

100 ディスク制御装置

110 チャネルアダプタ(チャネルアダプタ部)

111 データバッファ

120 キャッシュメモリ

130 共有メモリ

140 ディスクアダプタ(ディスクアダプタ部)

141 データバッファ

(6)

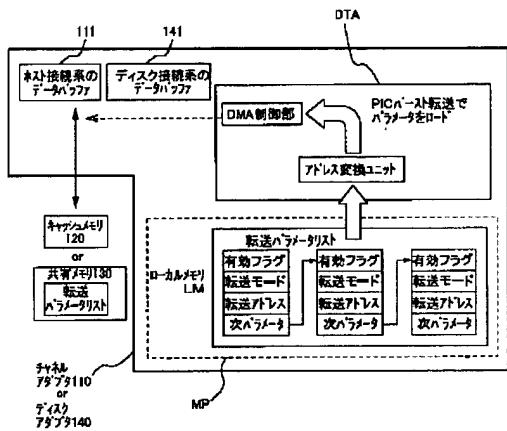
特開2003-345515

10

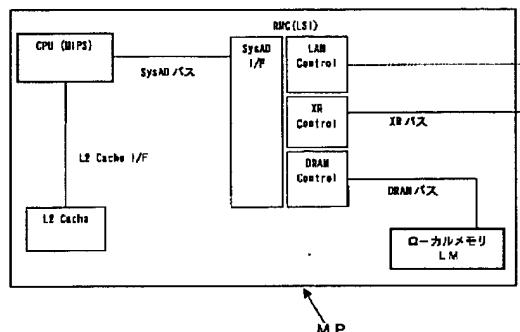
200 ディスクアレイ  
210 ディスク装置  
300 ホスト  
MP マイクロプロセッサ (プロセッサ)

\* LM ローカルメモリ (メモリ)  
DTA データ転送アダプタ (データ転送アダプタ手段)

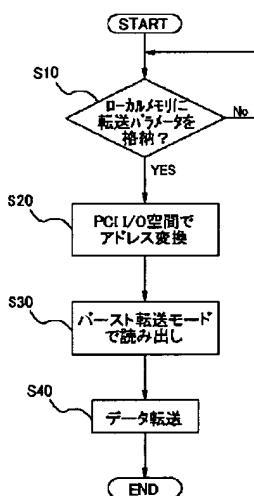
〔図1〕



[図2]



〔図3〕



【図4】

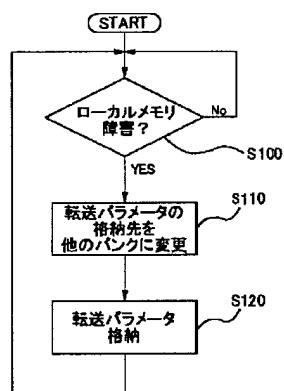
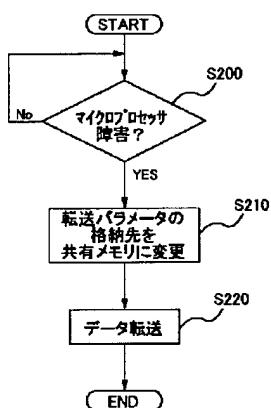
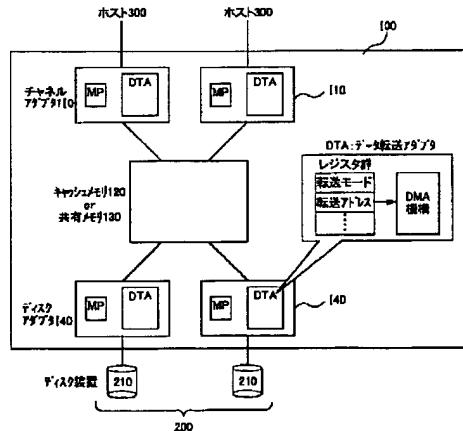


図51



[図6]



(7)

特開2003-345515

フロントページの続き

F ターム(参考) 5B014 EB03 EB05 GC01  
5B018 GA06 KA15 MA05  
5B061 DD01  
5B065 BA01 CA02 CA15 CE01 CE12  
CH01